Quantifying Seasonal Variation in Cloud Cover with Predictive Models

Ashok N. Srivastava, Ph.D. ashok@email.arc.nasa.gov
Deputy Area Lead, Discovery and Systems Health
Group Leader, Intelligent Data Understanding
NASA Ames Research Center

Joint Work with Rama Nemani, Ph.D. (NASA Ames)
Other Collaborators: Nikunj Oza, Ph.D., Mike Way, Ph.D., Jeff Scargle, Ph.D. (NASA Ames), Julienne Stroeve, Ph.D. (NSIDC, University of Colorado)

Outline of Talk

- Motivating problem*
- General Virtual Sensors problem
- Results
- Related accomplishments
- Summary

* Nearly all of this work was presented at the 2005 AGU conference.

Motivating Problem

Many remote sensing problems data analysis problems can be broken down into two components:

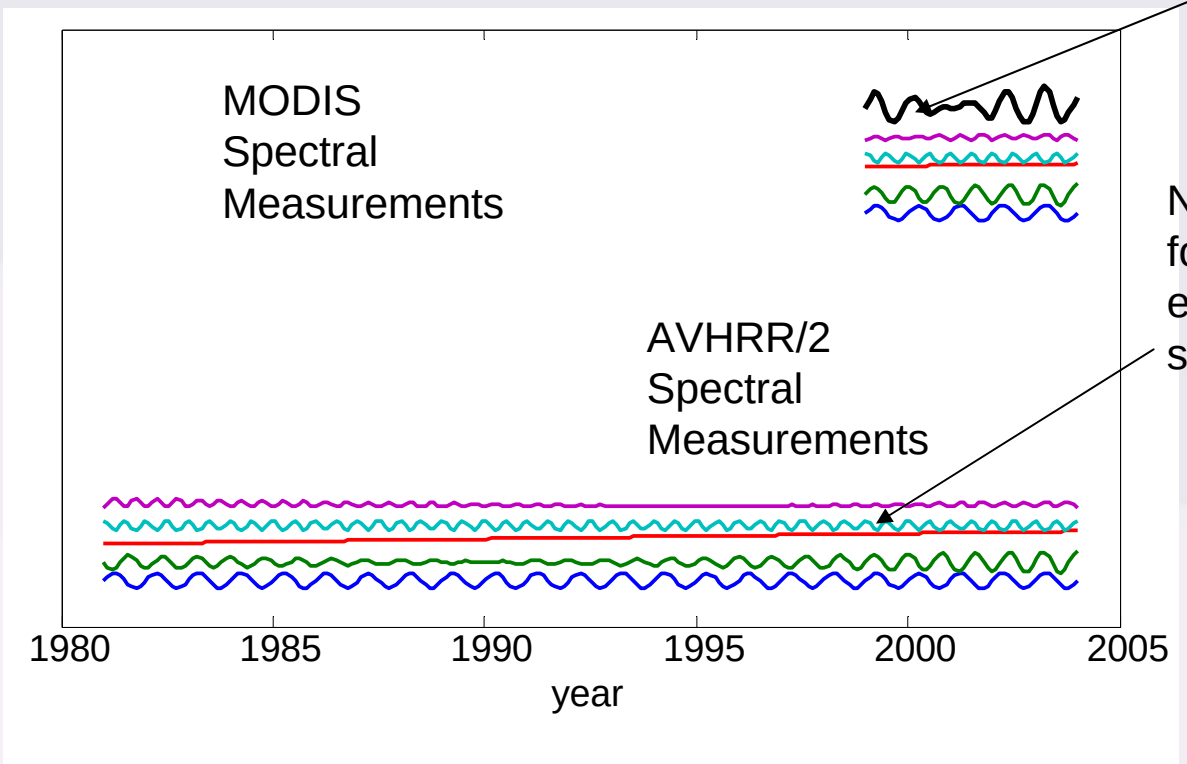
- (1) Seasonal variation
- (2) Variation induced by the model class

- The purpose of this study is to develop algorithms that help us identify these variations separately.

- It is necessary to understand this variation in order to develop stable models that can predict energy in one spectral band based on the energy in other spectral bands.

- We address these problems through the development of a Virtual Sensor.

Emulating Sensor Signals back in Time for Cloud Trending

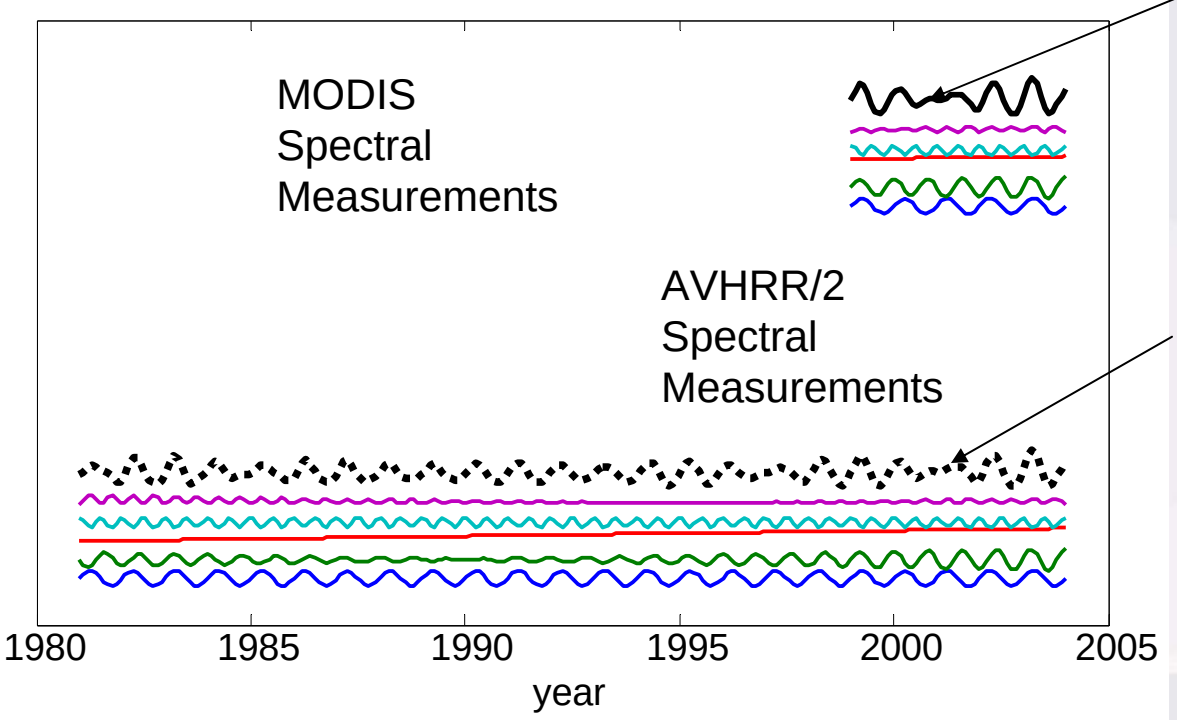


MODIS
1.6 micron Channel 6
(Black)

No Black Signal
for AVHRR/2. Need to
emulate AVHRR/2 cloud
signal

Cloud Detection back in Time

Solution: Predict 1.6μm channel using a Virtual Sensor



MODIS 1.6 micron Signal (Black)

Black Signal for AVHRR/2 estimated by a Virtual Sensor.

Multi Resolution Analysis

The Virtual Sensors concept can also be used to deal with multi-resolution analysis:

MODIS Channels 1 & 2: 250 m resolution
MODIS Channel 6: 500 m resolution.

Note: Channel 6 (at 1.6 microns) is not available at the 250 m resolution.

Virtual Sensors Approach

Given:

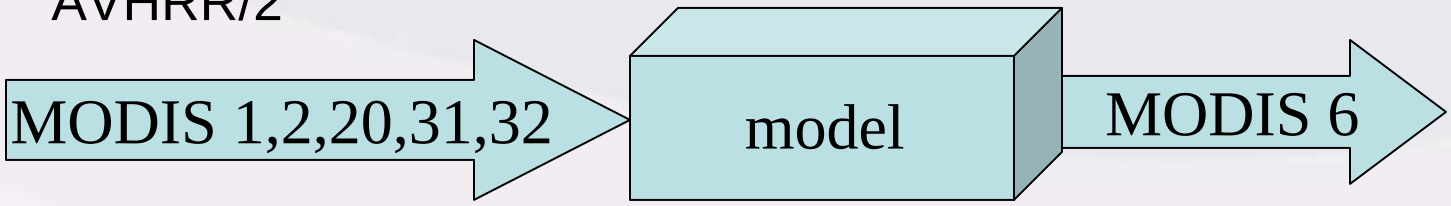
- MODIS channels 1, 2, 20, 31, 32 correspond to five AVHRR/2 channels

Develop:

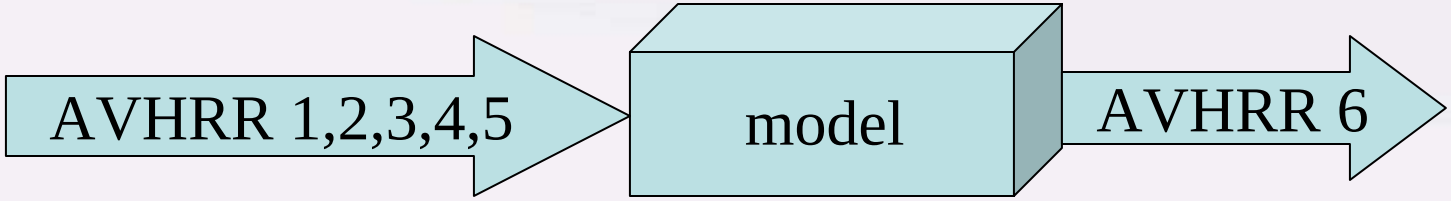
- Model MODIS channel 6 (1.6μm) as a function of five MODIS channels

Apply:

- Use function to construct estimate of 1.6μm channel for AVHRR/2



Model Construction

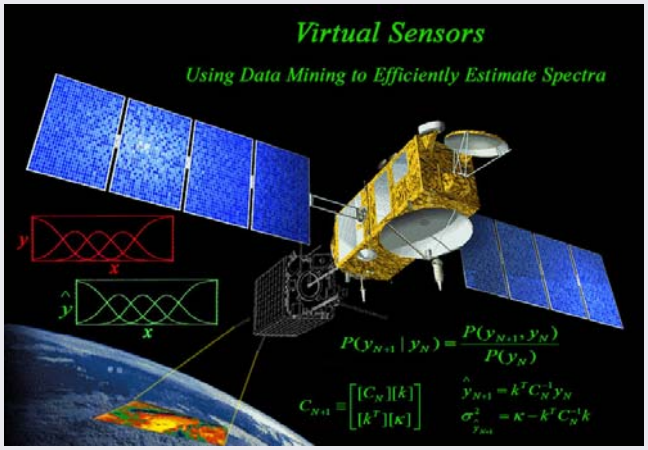


Model Application

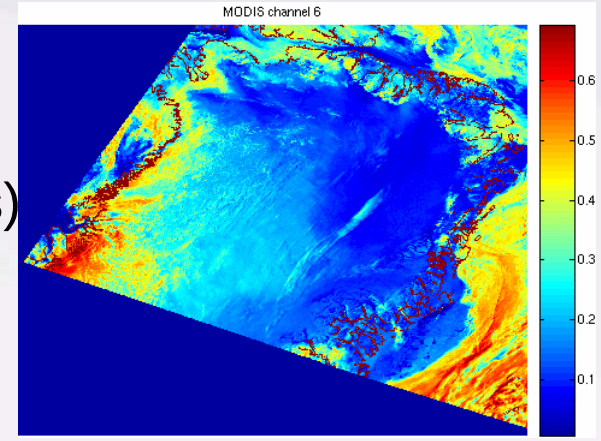
Virtual Sensors

Virtual Sensors predict the historical record for spectral measurements using relationships found from existing sensors and inputs from historical record.

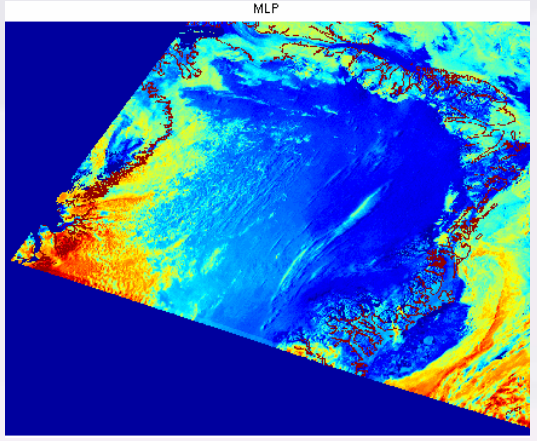
Useful for simulating sensors back in time or multi resolution analyses.



Accuracy of learned models for MODIS data: 70%-90% (over 2 weeks)



Channel 6



MLP prediction

Model Classes

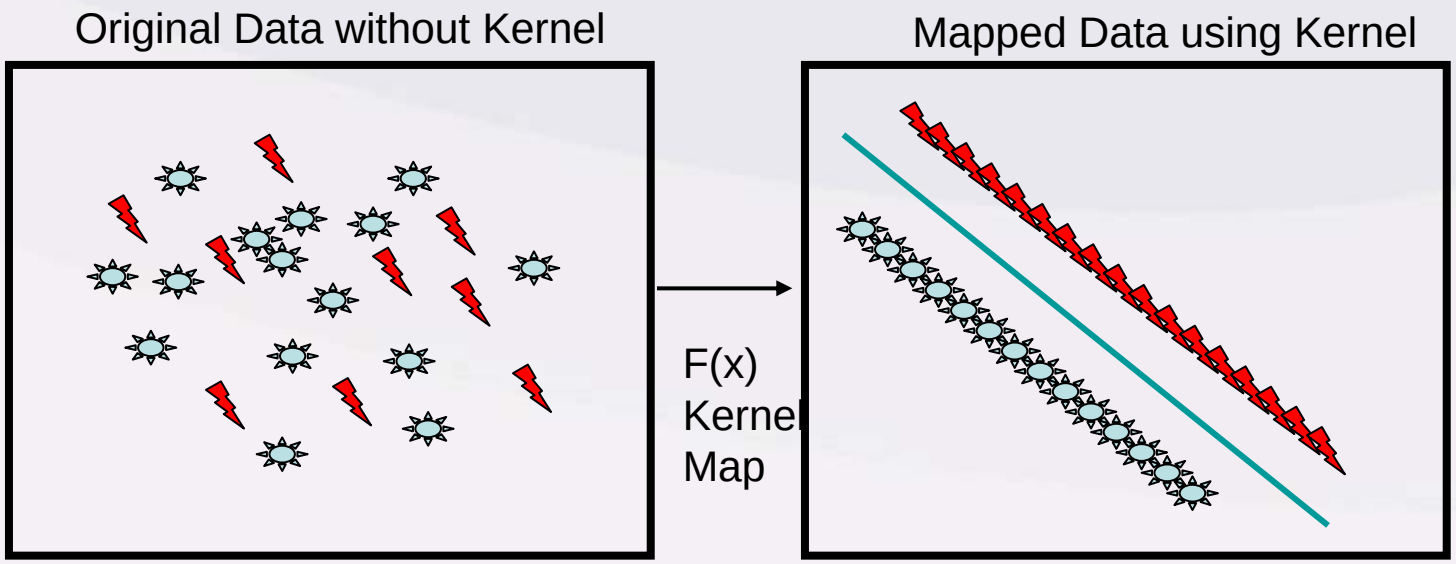
Linear Models:

Least squares regression (used as a very simple baseline)

Nonlinear Models:

Neural nets (used as a simple baseline)

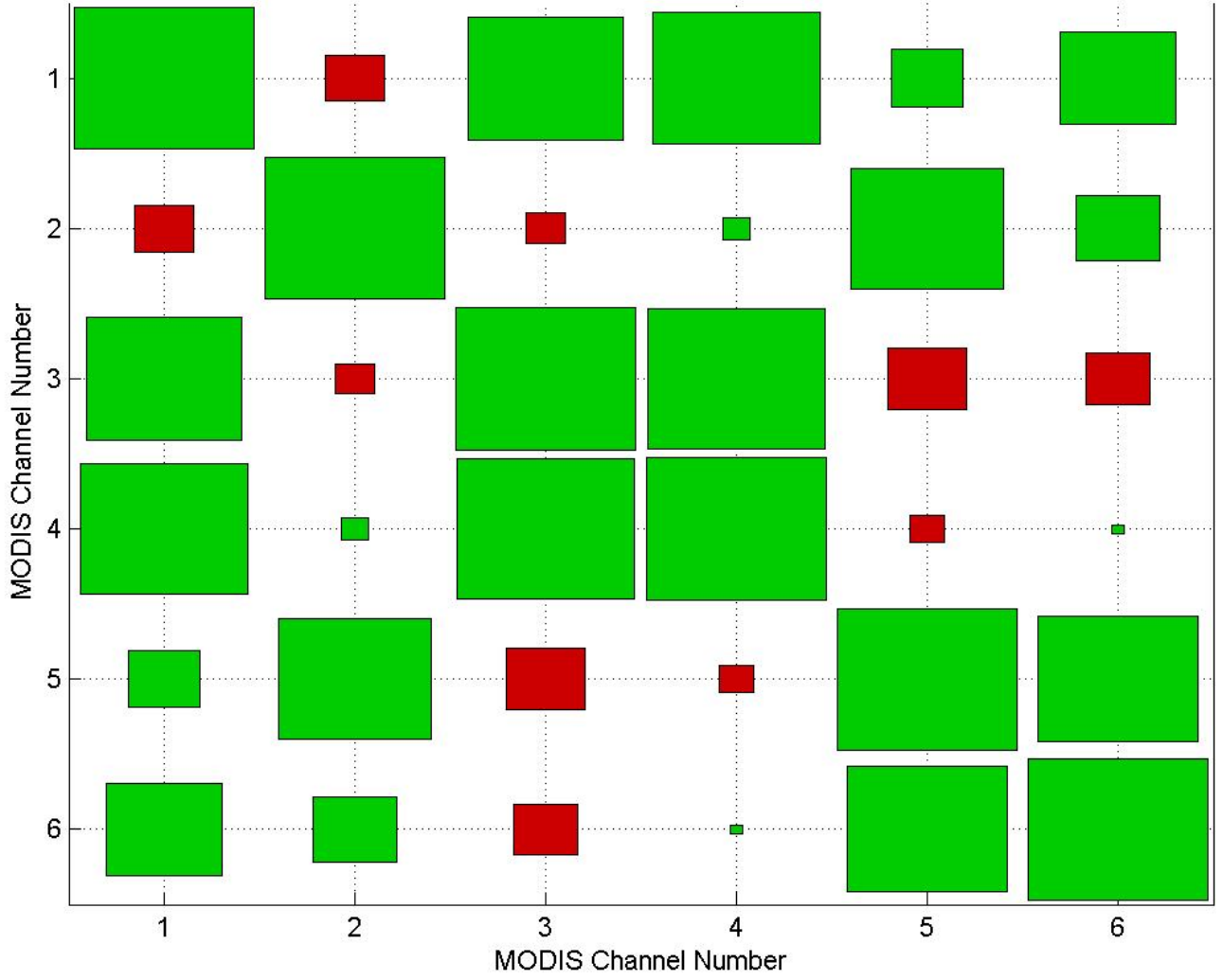
Gaussian Processes & Kernel Methods



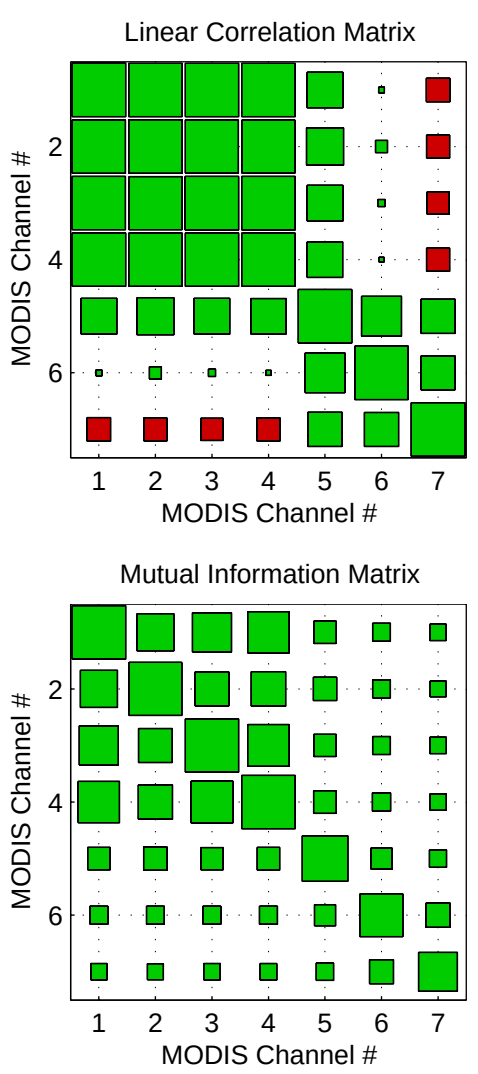
Data in original space: highly complex decision boundaries.

Data in high dimensional feature space can yield simple decision boundaries.

Linear Correlation Matrix for MODIS Channels over Fresno CA in 2005

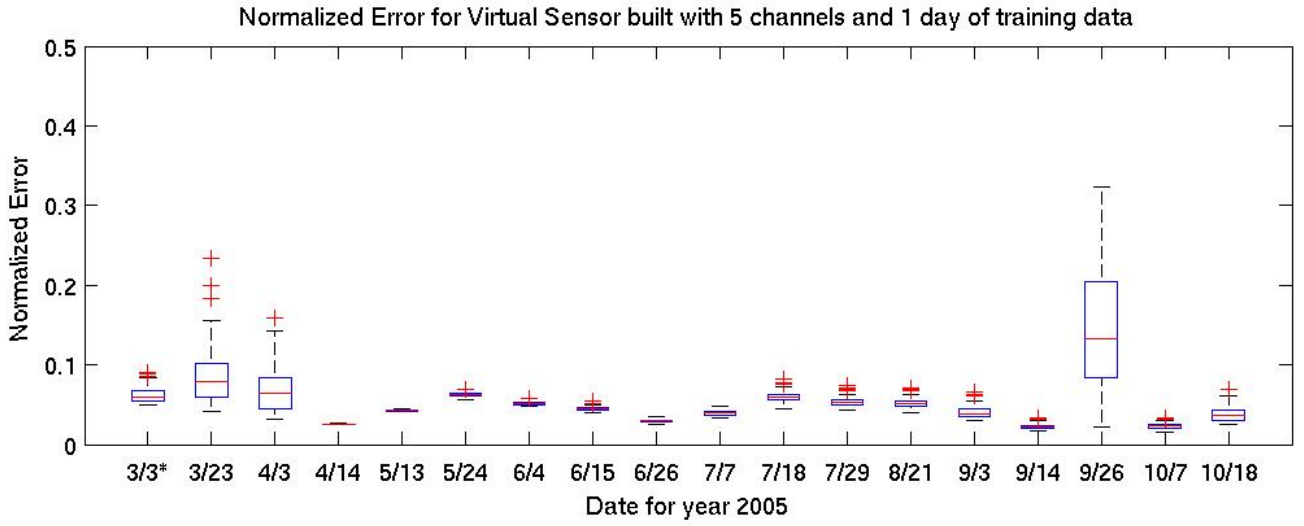


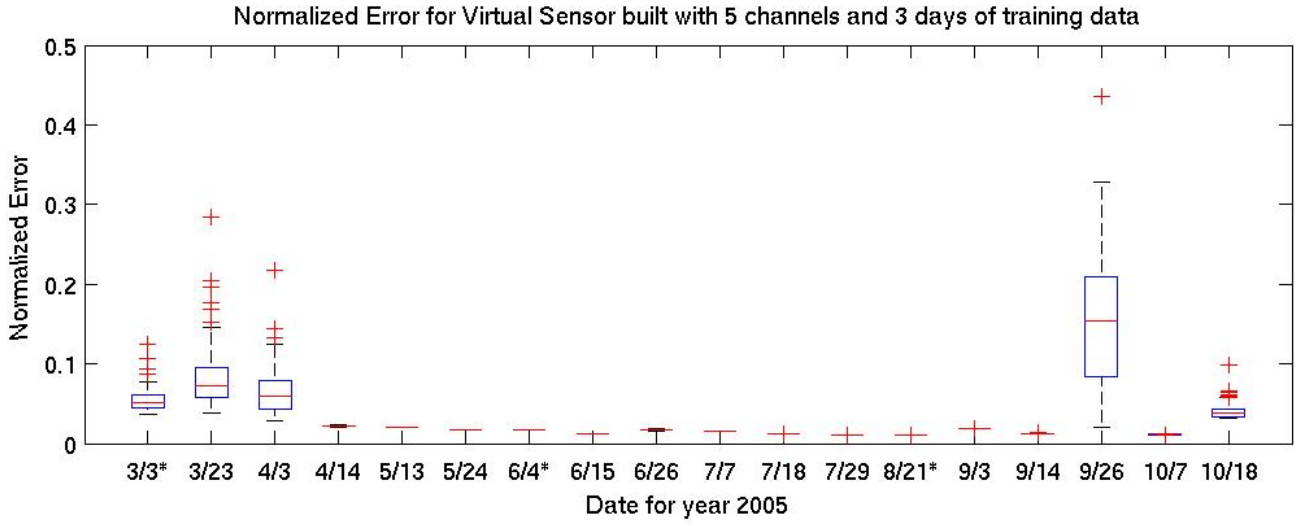
Correlation Matrix over Greenland, Year 2004



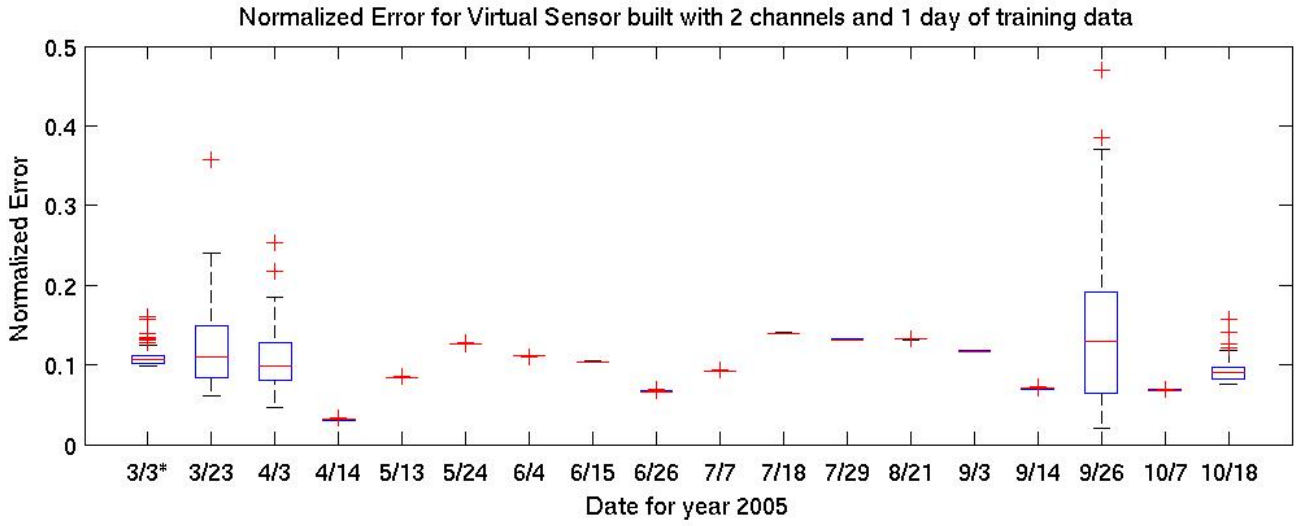
Correlation Matrices Changes with location

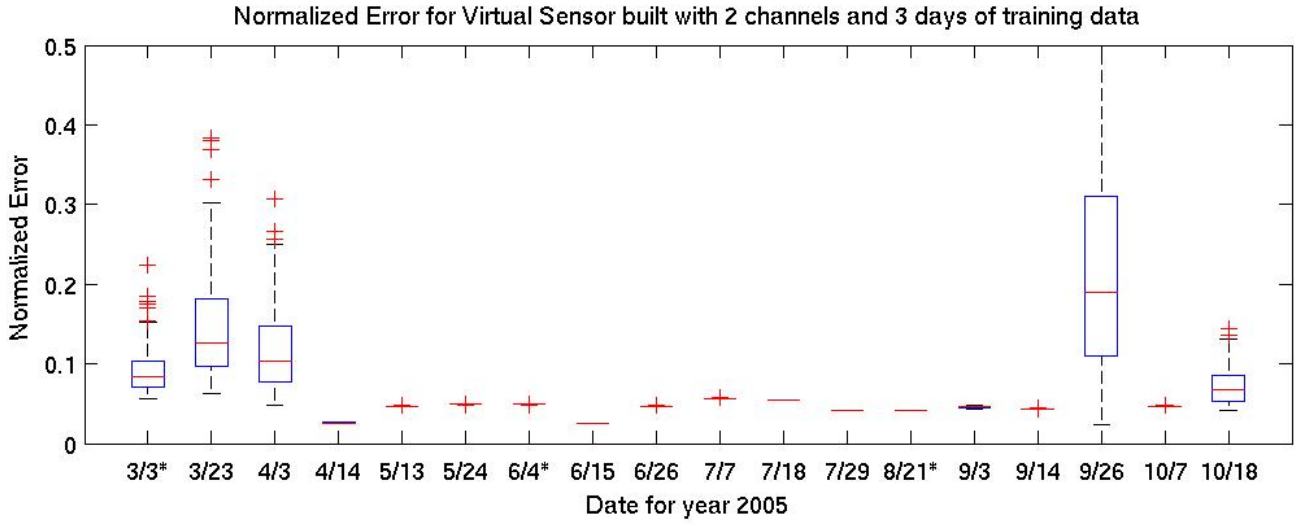
Seasonal Variation for 5 input channels



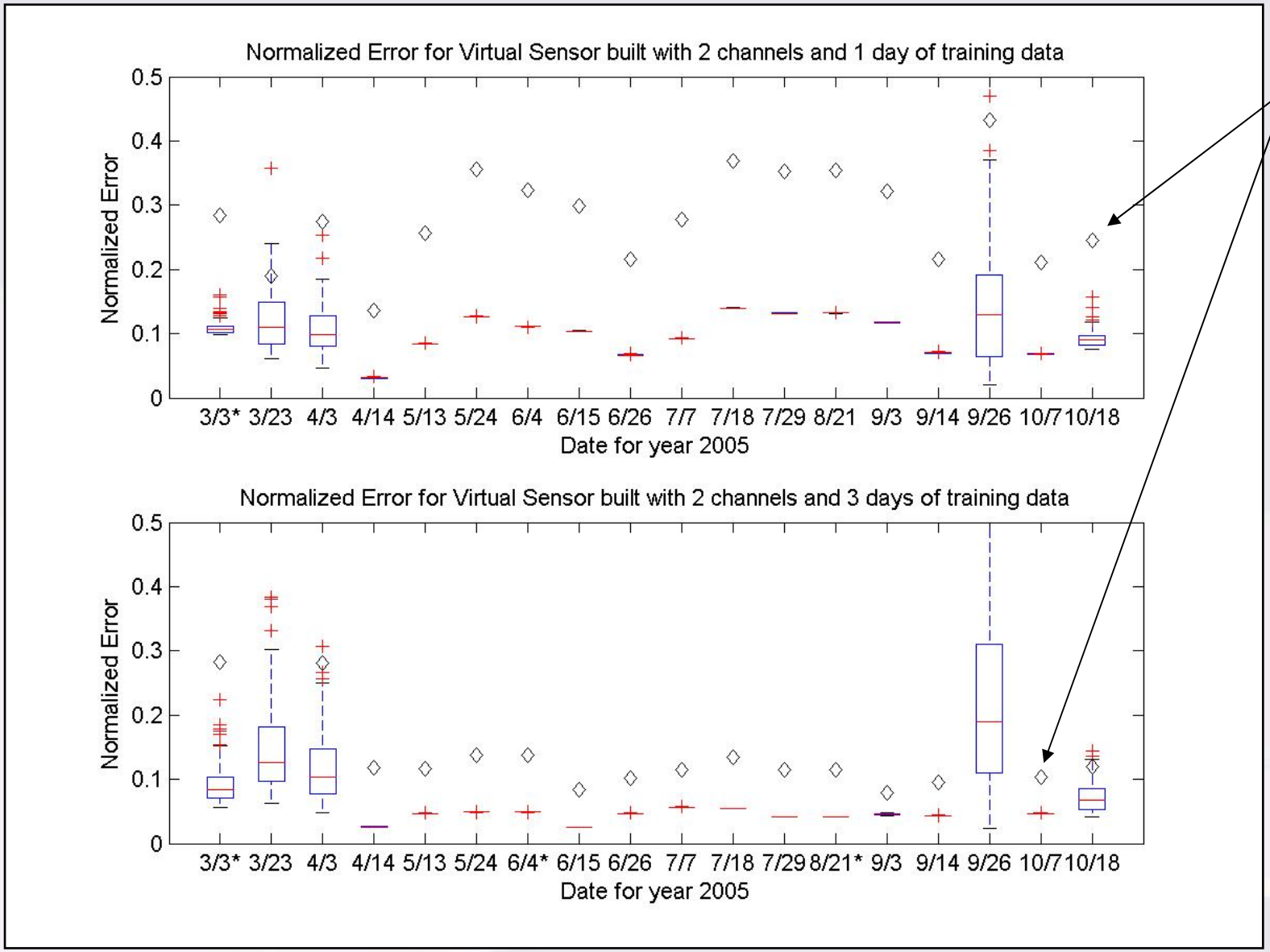


Seasonal Variation for 2 input channels



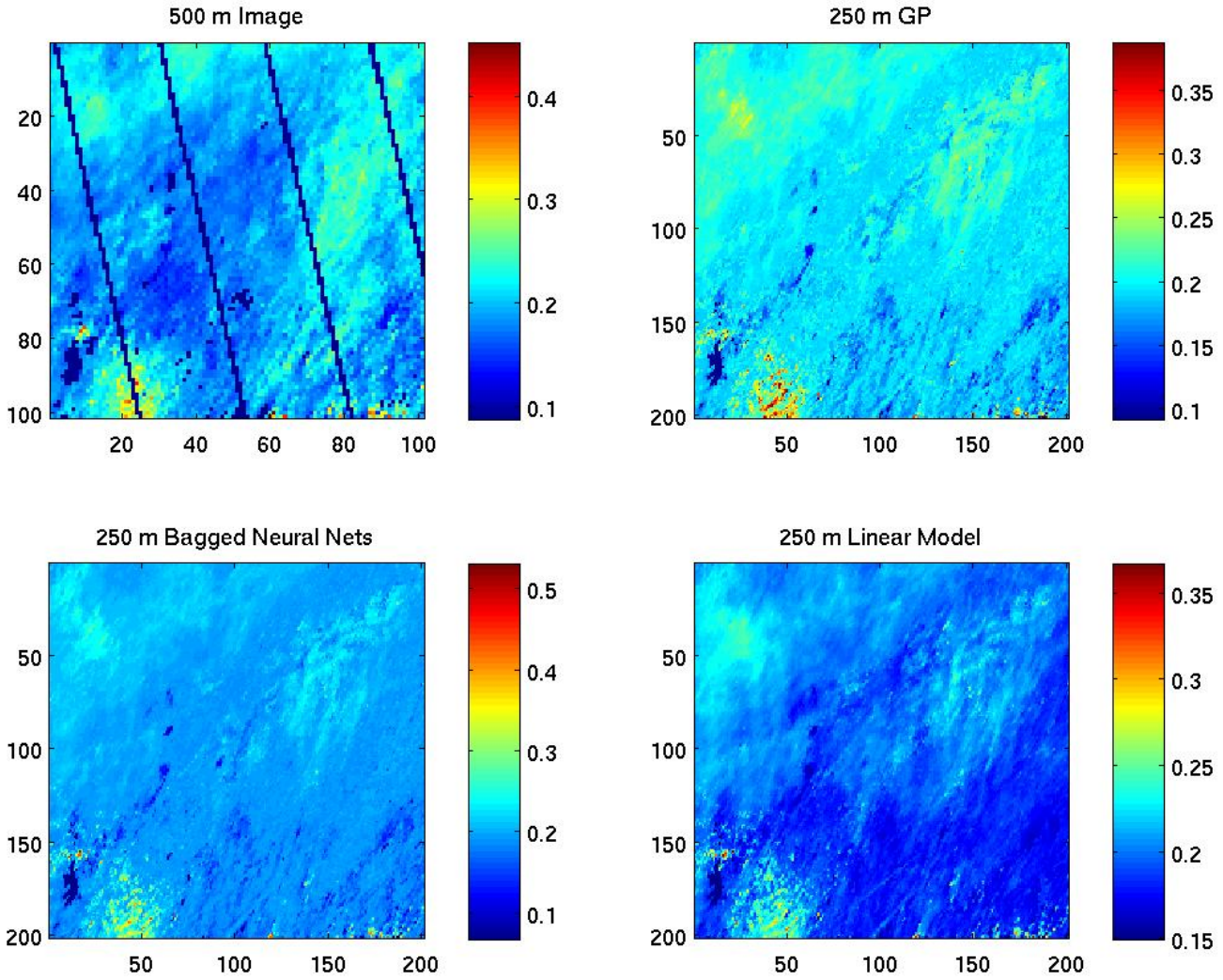


Seasonal Variation for 2 input channels

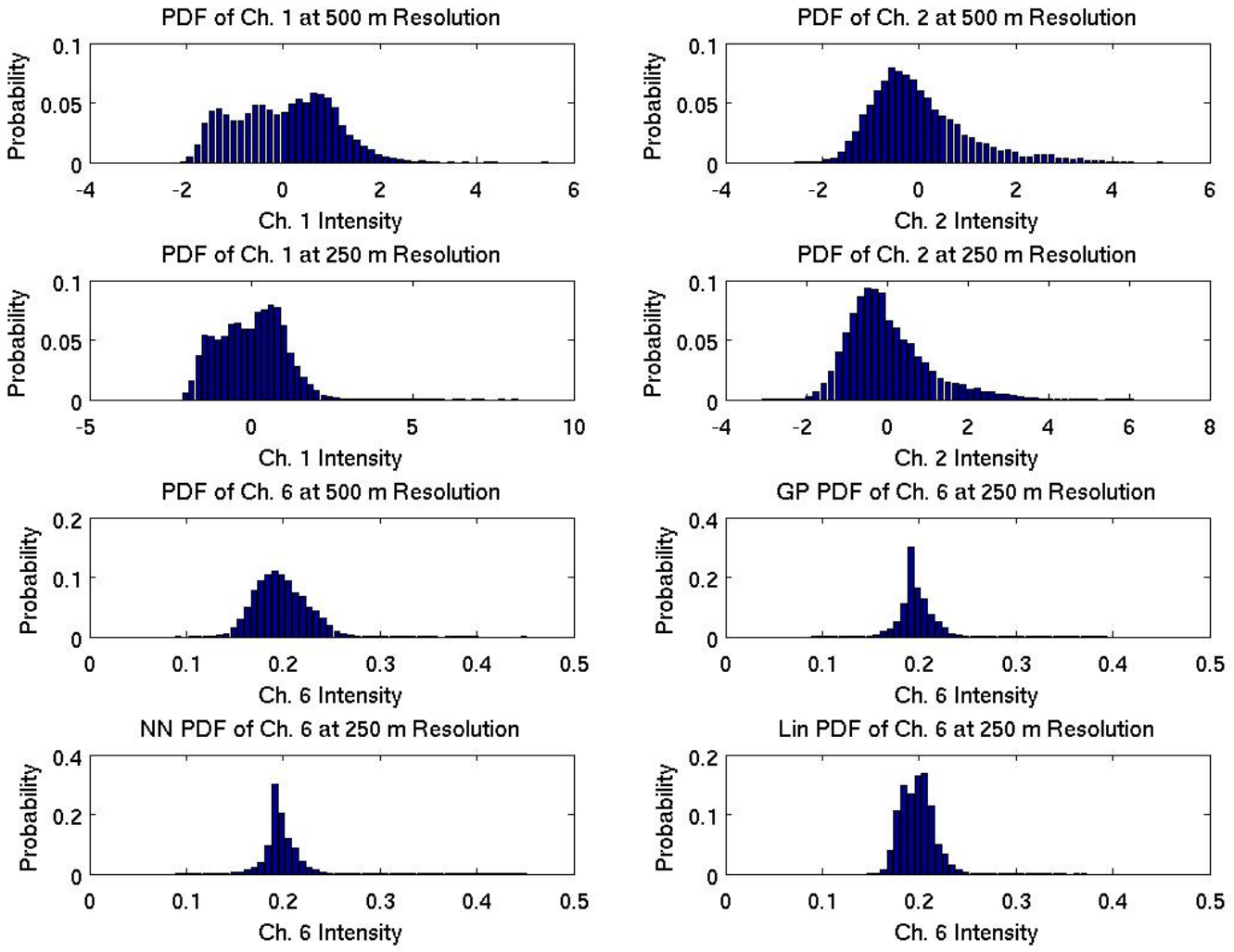


Black Diamonds: Linear Model

Multi-Resolution Predictions



Distributions of Multiresolution Predictions



Conclusions

The model class significantly affects the model's stability:

Linear Models: Produce higher overall error and are less robust to seasonal variation.

Nonlinear Models: Produce lower overall error and are more robust with respect to seasonal variations.

These results support the idea that a Virtual Sensor can be used to characterize sensor measurements through time or at different resolutions.

Can be used to reduce processing times significantly for some applications.